An Improved Force Field for Conformational Analysis of Sulfated Polysaccharides

DINO R. FERRO,* PAOLO PUMILIA, and MASSIMO RAGAZZI

Istituto di Chimica delle Macromolecole del C.N.R., via E. Bassini 15, I-20133 Milano, Italy

Received 11 July 1995; accepted 28 May 1996

ABSTRACT

A force field to be used in molecular mechanics studies of sulfated polysaccharides with explicit account of water and counterion interactions was derived from the analysis of six crystal structures of sulfated monosaccharide salts. The force field is based on Allinger's MM2, and was developed starting from the parameters used in previous studies of heparin and related oligosaccharides. While the novel parameters have been derived empirically, use of the atomic charge distribution obtained from *ab initio* quantum-mechanical computations, at the 6-31 + G** level, improves the quality of structural fitting significantly. The overall discrepancy between the positions of the nonhydrogen atoms determined by X-ray diffractometry and those corresponding to the minimum-energy structure is 0.21 Å. While most geometrical features of both carbohydrate and sulfate moieties are reproduced satisfactorily, in some cases (particularly in the case of the Na^+ salt of α-methyl-4-*O*-sulfogalactopyranoside) the hydrogen bond pattern is altered by energy minimization, probably due to errors in the balance of the strong electrostatic forces.

Preliminary results were presented at the Third International Satellite Symposium ''Conformational Analysis on Carbohydrates,'' Val Morin, Canada, 24–27 July 1994, abstract, p. 83. See also ref. 22.

* Author to whom all correspondence should be addressed.

CCC 0192-8651 / 97 / 030351-17

Introduction

Although sulfated polysaccharides, notably among them glycosaminoglycans (GAGs), form a class of macromolecules endowed with important biological functions,[1] the current treatment of their conformational characteristics by means of molecular mechanics is still far from satisfactory: major problems arise form their polyelectrolytic nature and, more specifically, from the presence of the groups $-O-SO_3^-$ and $-NH-SO_3^-$. To the best of our knowledge, no work of molecular mechanics or dynamics has so far attempted a realistic description of these polymers, with explicit inclusion of solvent and counterions in the molecular system. The main reason for this seems to be the lack of a reliable force field.

As a contribution to move one step forward with respect to the level of conformational analysis on GAGs carried in our laboratory in the last decade,[2–7] we present an improved force field based on semiempirical and *ab initio* computations on model compounds and on the crystal structure analysis of several sulfated monosaccharides. At the time when the first all-atom force field for GAGs was defined,[2,3] only one crystal structure of sulfated simple sugars was known (octa-sulfated sucrose[8]); therefore, that parameterization relied largely on structural data pertaining to various less suitable sulfated model compounds. Moreover, very crude approximations were assumed concerning the atomic charge distribution used to compute one of the major terms of the potential function, namely the electrostatic contribution. Since then, a growing number of crystal structures of sulfated derivatives of monosaccharides have been determined,[9–18] mainly due to the work of Mackie and Lamba.[9–13] The wealth of data gathered from these X-ray structure analyses recently led Lamba et al.[10] to propose a parameterization of the bending and stretching terms of the force field for *O*-sulfated monosaccharides, within the MM3 framework, as well as to analyze the details of cation coordination, crystal packing, and hydrogen-bonding pattern observed in such structures. On the other hand, Withfield and Tang[19] used accurate *ab initio* quantum-mechanical computations on methyl and ethyl sulfate anions to derive stretching, bending, and torsional parameters, as well as bond dipoles, for the MM2 force field; moreover, they analyzed the favorable H-bond arrangements of sulfate–water complexes and provided a theoretical insight into the electron charge distribution in those model compounds. Yet there is a need for defining a consistent force field including, besides the bonded parameters, all nonbonded interactions involving the sulfate (and other charged groups), carbohydrate, water, and counterion moieties, and for testing the validity of the parameters against the observed three-dimensional structures. Finally, after this work was completed, we became aware of a very recent article by Huige and Altona[20] on a force-field parameterization for sulfate and sulfamates based on *ab initio* calculations, leading to extensions of the AMBER and CHARMm fields; the implications of their results will also be discussed.

Methods

The force field previously utilized in the study of heparin and related sulfated oligosaccharides[2,3] was taken as the starting point for our attempt to obtain a force field which ultimately should allow for a realistic treatment of glycosaminoglycans in solution; i.e., with explicit inclusion of solvent and counterion interactions. For this purpose, we utilized five X-ray crystal structures of Na^+ and K^+ salts of *O*-sulfated monosaccharides, determined in the recent years by Lamba and coworkers.[9,10] A sixth structure, the Na^+ salt of dihydrated 2-*N*-sulfated glucosamine,[11] was also considered in the last stage of the work, mainly to check the transferability of the results. The only criterion adopted to determine the force-field parameters was the agreement between the experimental positions of the nonhydrogen atoms and the corresponding values calculated by energy minimization in Cartesian coordinates at constant unit cell. The adjustments of several of the parameters defining the force field were mainly done by trial-and-error, because an automatic least-squares procedure, used in the early stage of the refinement, did not converge to acceptable results.

The force field is essentially an extension of Allinger's MM2;[21] hence, we refer to the original definition of the potential functions and to the original MM2(87) parameters, apart from the specific differences stated below.

PARAMETERIZATION OF FORCE FIELD

(i) *Atomic classes*. Two new classes were defined for ions Na^+ and K^+. Class 7 (the same as for O=) was assigned to all the oxygen atoms of the $-O-SO_3^-$ and $-NH-SO_3^-$ groups; thus, no lone pair pseudo-atoms were assumed on these Os; the choice of dropping the lone pair from the oxygen bonded to the sugar ring, different from our previous one, brought a significant improvement in the structural fitting. The lone pair was kept on the N atom just for the purpose of using the previous bending parameters at N, but it was ignored as far as nonbonded interactions are concerned. For the water molecule we simply adopted a model based on Allinger's[21] definition of alcoholic O—H, thus assigning classes 6, 21, and 20 of MM2(87) to O_W, lone pair, and H_W, respectively.

(ii) *Stretching and bending parameters.* The entire set of parameters concerning the $C-O-SO_3^-$ moiety has been modified with respect to our early force field. In most of the observed structures of the model compounds one of the O^- atoms is approximately in the *anti* conformation relative to the vicinal carbon atom. Stereoelectronic effects arising from the coplanarity of the $C-O-S-O_a^-$ lead to a significant asymmetry of the SO_3^- group, whereby the bond angle $O-S-O_a^-$ is smaller than $O-S-O_g^-$ and the angle $O_g^- -S-O_g^-$ is larger than $O_g^- -S-O_a^-$. The asymmetry is usually not removed by deformations due to crystal packing. Although this situation is well reproduced by accurate quantum-mechanical computations,[19,22] it was soon found[2] that it cannot be simulated by molecular mechanics without distinguishing between *anti* and *gauche* O^-. Although this difficulty initially was, to a great extent, due to the electrostatic interactions between O^- and the lone pairs on O, we found that, even after dropping the lone pairs, the geometry of the sulfate group can be accurately reproduced only by adopting different θ_0 values for the mentioned bond angles. The parameters derived empirically here could, in the future, be refined in light of the results of *ab initio* computations.[19] As for the $-C-NH-SO_3^-$ moiety, we have temporarily: (a) adopted the standard MM2(87) parameters for C—N, N—H, N—Lp, C—N—H, C—N—Lp, and H—N—Lp; (b) used the previous parameters[3] for N—S, C—N—S, and S—N—Lp; and (c) used the same values found for $O-SO_3^-$ for the other parameters of $N-SO_3^-$.

All other bending and stretching parameters were taken from MM2(87), but to account for the different treatment of the electrostatic term, the values of θ_0 for bond angles O—C—O and C—O—C were increased by 4° and 2°, respectively. Finally, for the water bond angle $H_W-O_W-H_W$, not defined in MM2(87), we assumed the values of $\theta_0 = 103.3°$ and $K_\vartheta = 0.348\ \text{md} \cdot \text{Å} \cdot \text{rad}^{-2}$.

(iii) *Torsional barriers.* In the late stages of the refinement, it was found that the overall structural fitting improves by introducing a threefold torsional barrier for each dihedral angle $C-O-S-O^-$, with $V_3 = 0.4\ \text{kcal} \cdot \text{mol}^{-1}$. This empirical value has been kept, although it corresponds to a total barrier of $1.8\ \text{kcal} \cdot \text{mol}^{-1}$ for the rotation about O—S in $CH_3-O-SO_3^-$, against a value of $2.7\ \text{kcal} \cdot \text{mol}^{-1}$ computed *ab initio*.[19,22] The same value of V_3 has also been used for the dihedrals $C-N-S-O^-$ and $H-N-S-O^-$, while, as in the past, no torsional term has been considered for the rotations about the C—O and C—N bonds connected to the sulfate groups, because no empirical indication was suggested by the crystal structure simulations.

The final set of novel bonded parameters are listed in Table I.

(iv) *van der Waals parameters.* A unique set of van der Waals potential functions was finally adopted for treating intermolecular and intramolecular nonbonded interactions, including all first-neighbor $(1 \cdots 4)$ terms, although in the intermediate stage of the work it seemed necessary to introduce some distinctions between the three types of interactions. Both Buckingham and Lennard–Jones potential functions were utilized, hence the results will be presented in Table II in the more general form:

$$U_{ij} = -Ar_{ij}^{-6} + Br_{ij}^{-n}\exp(-\beta r_{ij})$$

All parameters for $Na^+ \cdots X$ and $K^+ \cdots X$ interactions were defined anew and adjusted independently to minimize the crystal structural discrepancies. Several of these parameters, concerning atoms weakly interacting with the ions, are still to be considered as ill-determined. All other parameters are taken from MM2(87), except for those concerning the interaction between hydrogen and lone pair, which were previously modified[1] ($r^* = 1.6$ Å, $\varepsilon^* = 0.048\ \text{kcal} \cdot \text{mol}^{-1}$) to better account for hydrogen bonding.

(v) *Hydrogen-bond parameters.* Special 6-12 or 6-*exp* functions are used to treat the interactions between hydrogen and acceptor when a hydrogen bond is formed. The parameters of these functions have been revised with respect to our previous force field[2,3] (see bottom of Table II).

(vi) *Electrostatic interactions.* The electrostatic energy contribution is computed in the monopole approximation:

$$E_{el} = \sum_{i,j} Q_i \cdot Q_j / \varepsilon \cdot r_{ij}$$

where the sum is extended to all nonbonded intramolecular atom pairs and to the intermolecular pairs within a specified cut-off distance. The cut-off was applied so that if a given atom pair is included, then all the interactions between the two (neutral) molecules containing the two atoms are also included: the chosen value of 11 Å seems sufficient to obtain convergence of the minimum-energy structure, if not of the absolute value of the total electrostatic energy.

For the dielectric constant, ε, while in calculations simulating the conditions of a macromolecule in solution, we usually adopted an effective value of 3, in the present crystal simulation we use the value of 1. This choice seems consistent with the facts that no polarization effects are considered in our model and that, having explicitly included the interactions with the molecules surrounding the asymmetric unit up to a sufficient distance, the *vacuum* is the proper medium for our system. However, a serious problem is encountered in the computation of the *intramolecular* electrostatic energy: since we omit the interactions between atoms separated by less than three bonds, the contribution of the charges on a pair of nearest-neighbor atoms ($1 \cdots 4$) arising from the dipoles of the connecting bonds 1—2 and 3—4 are not balanced by the interactions between bonded atoms. Such an overestimate of the $1 \cdots 4$ Coulomb interactions (e.g., in comparison with the point dipole $\cdots$ dipole treatment) may cause significant bond-angle distortions where large dipoles are involved (e.g. at the anomeric site), and would require a different parametrization of the bending parameters when using different values of ε. To bypass this difficulty, we first chose to use $\varepsilon = 3$ for all the intramolecular interactions even in the crystal calculations. Although this choice underestimates the interactions between atoms separated by several bonds, it has the advantage of avoiding a discontinuity when $\varepsilon > 1$ is used to simulate the screening effect of solvent. Moreover, previous computations on simple model compounds (e.g., dimetoxymethane) had shown that the original MM2 parameterization, based on point dipoles and $\varepsilon = 1.5$, can be adequately maintained (with adjustments of some bending parameters, as previously mentioned) in the monopole approximation if we use the value of 3 for ε. However, our final choice

TABLE I.
Modified Stretching, Bending, and Torsional Parameters of MM2 Force Field for C—O—SO_3^- and C—NH—SO_3^- Groups.

Bond	K_s	l_0	l_{0corr}
C—O_s	5.36	1.444	
O_s—S	5.00	1.60	
S—O^-	8.4	1.446	
C—N	5.1	1.438	1.453[a]
N—S	3.0	1.65	
N—H	6.1	1.02	

Bond angle	K_θ	θ_0
C—O_s—S	0.7	109.0
O_s—S—O_a^-	1.4	101.0
O_s—S—O_g^-	1.4	104.0
O_g—S—O_a^-	1.4	114.0
O_g—S—$O_{g'}^-$	1.4	113.0
C—N—S	0.7	112.0
C—N—H	0.5	109.47
S—N—H	0.5	111.0
C—N—Lp	0.5	109.2
H—N—Lp	0.5	108.0
S—N—Lp	0.5	109.2
N—S—O_a^-	1.4	101.0
N—S—O_g^-	1.4	104.0

Other bond angles	K_θ	θ_0
O—C(HR)—O	0.46	101.0
C—O—C	0.77	108.8
H_w—O_w—H_w	0.348	103.3

	Torsional barriers		
Dihedral	V1	V2	V3
C—O—S—O^-	0.0	0.0	0.4
C—N—S—O^-	0.0	0.0	0.4
H—N—S—O^-	0.0	0.0	0.4
X—C—O—S	0.0	0.0	0.0
X—C—N—S	0.0	0.0	0.0
X—C—N(S)—H	0.0	0.0	0.0

[a] After electronegativity correction.

TABLE II.
van der Waals Parameters for Ions Na^+ and K^+ Interacting With Sulfated Carbohydrates, and Hydrogen-Bond Parameters, According to the General Expression: $U_{ij} = -Ar_{ij}^{-6} + Br_{ij}^{-n}\exp(-\beta r_{ij})$.

Class	Atom	Ion	A	B / 1000	β	n
1	Csp^3	Na^+	129.40532	26.361	4.26621	0
2	Csp^2	Na^+	101.26420	39.440	4.75285	0
3	C′	Na^+	120.47843	33.640	4.49640	0
5	HC	Na^+	36.00480	26.187	5.27426	0
6	O	Na^+	490.	45.0	3.880	0
7	O^-	Na^+	250.	120.0	4.165	0
8	N	Na^+	100.	400.0		12
16	S	Na^+	400.	2000.0		12
20	Lp	Na^+	16.	.5	5.05225	0
21	HO	Na^+	36.00480	26.187	5.27426	0
23	HN	Na^+	36.00480	26.187	5.27426	0
1	Csp^3	K^+	131.13758	60.9	4.89428	0
2	Csp^2	K^+	402.19918	202.42	4.96032	0
3	C′	K^+	400.41895	160.66	4.77646	0
5	HC	K^+	82.73906	33.35	4.78011	0
6	O	K^+	412.67419	161.25	4.46621	0
7	O^-	K^+	1260.	225.0	3.900	0
8	N	K^+	321.70613	74.82	4.36148	0
16	S	K^+	1921.64636	158.34	3.66882	0
20	Lp	K^+	240.	12.0	3.53501	0
21	HO	K^+	83.00473	28.681	5.35780	0
Acceptor	Donor					
O	HO		13.393	23.2	5.430	
O^-	HO		13.393	46.4	5.430	
N	HO		17.000	20.0	5.211	
O^-	HN		10.000	46.4	5.430	

has been that of using the same value of ε (1 in crystal calculations) for both *intra*- and *inter* molecular interactions, but scaling the charges due to the dipoles of the connecting bonds in the case of $1\cdots4$ interactions; that is, the scaling factor of 0.1, taken empirically, yields slightly better results than the previous choice.

The charge distribution of the carbohydrate moiety (and the water molecule) is still obtained from the MM2(87) dipole moments. For the $-C-O-SO_3^-$ group, the empirical atomic charges used in the previous force field were adopted in the early stage of this work and soon replaced by a set of values derived from semiempirical quantum-mechanical (PM3) computations on methyl-*O*-sulfate.[22] The Merz–Kollman–Singh (MKS) scheme of fitting,[23] as implemented in MOPAC 6.0 (ref. 24), was used. These computations evidenced a serious flaw in the previous simplistic assumptions; in fact, while a null charge had been assumed on the sulfur atom, the quantum-mechanical calculations show a strong polarization of the S—O bonds, leaving much larger negative charges on the oxygens and a net charge of nearly $+2e$ on sulfur. Thus, the electrostatic interactions may have been seriously underestimated in the previous force field. The semiempirical PM3 charges were used during most of the work for parameterizing the new force field, until the overall rms displacements of nonhydrogen atoms reached the value of 0.26 Å.

A more accurate charge distribution was then obtained from *ab initio* quantum-mechanical computations on methyl-*O*-sulfate, performed with Gaussian 92 (ref. 25) at the 6-31 + G** level with the post-SCF Møller–Plesset electronic correlation treatment at the second order (MP2); charges were evaluated with MKS fitting, using the DENSITY = MP2 keyword. Other details of these computations are presented elsewhere.[26] Substitution of the PM3 charges with the *ab initio* MP2 set, without any other change in the force-field parameters, lowered the overall discrepancy significantly. Thus, the *ab initio* charge distribution was adopted for the final refinement of the force field, both on the grounds of theoretical considerations and of the agreement

with the experimental structural data. The different sets of atomic charges used for the $-C-O-SO_3^-$ and $-C-NH-SO_3^-$ moieties are given in Table III.

COMPUTER PROGRAM FOR MOLECULAR MECHANICS

All molecular modeling, as well as the fitting of the force-field parameters, was carried out by means of the program for molecular mechanics, CHAMP,[27] developed in our laboratory. Among many special features provided by this program, of particular relevance in this work was its ability to perform energy minimization of (macro)molecular systems in the presence of crystal (and/or helical) symmetry, including cases with atoms in special positions (e.g., crystal **4**), with care of preserving electroneutrality of the microcrystal. A new version was generated to compute the intramolecular electrostatic energy by scaling the nearest-neighbor dipoles, as outlined above. All computations were done on a Silicon Graphics IRIS 4D/320GTX workstation.

Results and Discussion

Hydrogen bonding certainly plays an important role among the forces that determine the actual structure of the crystals under investigation in the present work. Unfortunately, the positions of hydrogen atoms, the polar ones included, are determined with low accuracy by X-ray diffractometry, so that in some cases their geometrical features depart significantly from acceptable values. Therefore, in a first step, the positions of all the hydrogen atoms and lone-pair pseudo-atoms of the five crystal structures **1**–**5** of *O*-sulfated monosaccharides (schematically shown in Fig. 1) were optimized starting from the experimental coordinates, keeping the heavy atoms fixed and using a preliminary force field.[2,3] Such an optimization was then repeated several times, as the force-field fitting improved. Moreover, alternative orientations of the hydroxyl groups and of the water molecules were tested to check whether H-bonding patterns different from those proposed on the basis of X-ray data analysis could be energetically favored or lead to a final structure (i.e., after all-atom minimization) in closer agreement with experiment. Indeed, this was the case with α-MeGal-4S K (crystal **4**): in the original structure, and also after hydrogen optimization, bond O_3-H points toward one of the O^- of a surrounding molecule, the $O_3 \cdots O^-$ distance being 3.25 Å. However, O_3 is closer to the O_5 of another molecule, so that rotation of O_3-H by $\sim 120°$ leads to the formation of a hydrogen bond $O_3-H \cdots O_5$ and to the decrease of the total energy by a few kilocalories per mole. This alternative H-bond pattern produced a decrease of the heavy-atom rms deviation of about 0.2 Å. We also carried out calculations on **4** where both O_3-H conformations were taken into account in equilibrium, with various statistical weights. The result was that the rmsd increases with the weight of the original conformation, which was then definitively dropped from subsequent computations.

For the other crystals, modifications of the hydrogen-bonding pattern occurred only during the

TABLE III.
Different Sets of Net Atomic Charges Used for Groups $S-O-SO_3^-$ and $C-NH-SO_3^-$.[a] Total Charge on Groups Is 1 e, a Null Charge Being Assigned to H Atoms Bonded to C, According to MM2 Scheme.

	$C-O-SO_3^-$			$C-NH-SO_3^-$	
		Ab initio		Ab Initio	
Atom	Semiempirical PM3	MP2	SCF	MP2	SCF
C	0.074	0.158	0.182	0.181	0.193
O/N	−0.525	−0.422	−0.446	−0.765	−0.810
H_N				0.376	0.393
S	1.931	1.144	1.441	1.143	1.423
O_a^{-} [b]	−0.800	−0.622	−0.723	−0.658	−0.743
O_g^{-} [c]	−0.840	−0.629	−0.727	−0.617	−0.706
$O_{g'}^{-}$	−0.840	−0.629	−0.727	−0.658	−0.750

[a] From ref. 23.
[b] In position *anti* to C.
[c] In position *anti* to H_N.

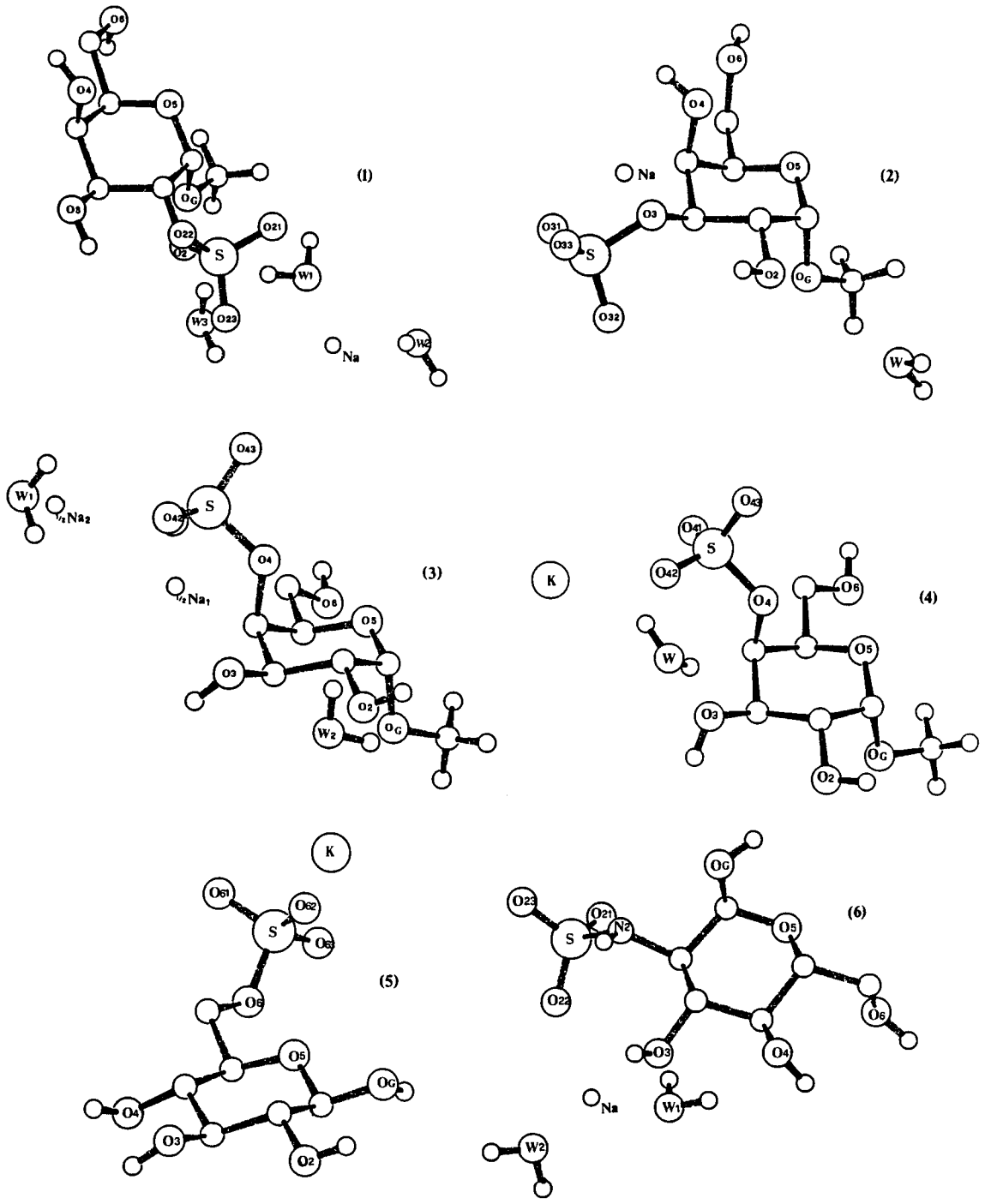

FIGURE 1. Schematic representation of the six sulfated monosaccharide salts: (**1**) α-MeGal-2S Na; (**2**) α-MeGal-3S Na; (**3**) α-MeGal-4S Na; (**4**) α-MeGal-4S K; (**5**) β-Glc-6S K; (**6**) α-Gal-2NS Na.

all-atom energy minimization: in such cases, particularly with crystals **1** and **5** which contain three and two water molecules per a.u., the different H-bond possibilities were considered in the subsequent stages of the force-field fitting.

The second step of the procedure consists in the crystal energy minimization with respect to the Cartesian coordinates of all atoms of the asymmetric unit, keeping the unit cell constant, followed by the analysis of the discrepancies with experiment,

TABLE IV.
Results of Energy Minimizations on Five Crystals of *O*-Sulfated Monosaccharides.

(a) Root-mean-square deviation between computed and experimental atomic coordinates (nonhydrogen atoms)

Crystal	α-MeGal-2S Na (1)	α-MeGal-3S Na (2)	α-MeGal-4S Na (3)	α-MeGal-4S K (4)	β-Glc-6S K (5)	Overall	Notes *charges*, fit
rms dev. (Å)	0.172	0.224	0.324	0.334	0.201	0.260	*Semiemp.*, fit-I
rms dev. (Å)	0.176	0.160	0.302	0.221	0.188	0.217	*Ab initio MP2*, fit-I
rms dev. (Å)	0.208	0.137	0.290	0.163	0.144	0.199	*Ab initio MP2*, fit-II
rms dev. (Å)	0.197	0.156	0.288	0.159	0.160	0.200	*Ab initio SCF*, fit-II

(b) Comparison between geometrical features of the *O*-sulfate group (Fit-II, with ab initio MP2 charges)

	α-MeGal-2S Na			α-MeGal-3S Na			α-MeGal-4S Na			α-MeGal-4S K			β-Glc-6S K			Overall	
Crystal	Obs.	Calcd.	Err.	Obs.	Calcd.	Err.	Obs.	Calcd.	Err.	Obs.	Calcd.	Err.	Obs.	Calcd.	Err.	rmsd	Av. dev.
Bond length																	
O_s—S	1.600	1.596	−.004	1.602	1.599	−.003	1.579	1.593	.014	1.587	1.589	.002	1.564	1.592	0.28	0.014	0.007
S—$O_{g/g'}$	1.413	1.441	.028	1.430	1.441	.011	1.445	1.442	−.003	1.433	1.441	.008	1.452	1.442	−.010	0.015	−0.007
S—O_a	1.435	1.442	.007	1.444	1.443	−.001	1.437	1.440	.003	1.457	1.439	−.018	1.414	1.441	.027	0.015	.004
Bond angle																	
C—O—S	115.9	116.6	0.7	118.8	117.8	−1.0	121.8	121.5	−0.3	120.1	120.7	0.6	118.9	118.1	−0.8	0.7	−0.2
O—S—O_g	105.2	106.1	0.9	106.9	106.3	−0.6	108.0	107.8	−0.2	108.1	108.0	−0.1	105.6	106.3	0.7	0.6	0.1
O—S—$O_{g'}$	106.2	105.5	−0.7	106.8	105.9	−0.9	107.8	105.1	−2.7	106.5	106.8	0.3	107.9	106.9	−1.0	1.4	−1.0
O—S—O_a	103.4	103.5	0.1	100.7	100.3	−0.4	101.8	101.7	−0.1	101.8	101.4	−0.4	102.7	101.8	−0.9	0.5	−0.3
O_g—S—$O_{g'}$	112.6	113.6	1.0	113.0	113.5	0.5	110.8	109.8	−1.0	112.6	112.9	0.3	111.7	113.0	1.3	0.9	0.4
O_a—S—O_g	113.9	113.0	−0.9	114.3	114.2	−0.1	114.4	117.1	2.7	112.0	112.9	0.9	112.4	113.6	1.2	1.4	0.8
O_a—S—$O_{g'}$	114.2	114.0	−0.2	113.8	115.0	1.2	113.3	114.2	0.9	114.8	113.8	−1.0	115.5	114.0	−1.5	1.0	0.1
Dihedral																	
C—C—O—S	−131.8	−121.9	−9.9	94.5	100.4	5.9	−122.1	−130.6	−8.5	−128.5	−123.6	4.9	−141.2	−132.5	8.7	7.8	
C—O—S—O_a	177.4	171.8	−5.6	176.8	174.8	−2.6	−160.8	−152.6	8.2	−169.2	−162.5	6.7	175.1	172.2	−2.9	5.6	

a gradual modification of the force-field parameters, a new minimization, and so on. As we mentioned previously, several alternative choices concerning various features of the force field were considered along this course. In principle, the crystal energy should be minimized also with respect to the unit cell parameters. However, past accurate crystal packing calculations on polymeric systems,[28] also carried out with MM2-like potentials, showed that, while energy minimization with respect to the cell parameters brings up a considerable shrinking of the unit cell, due to the neglect of thermal expansion effects, minimization at constant cell allows for significant structural refinements and in some instances clarifies the origin of conformational or structural modifications. Conversely, in the present work, free-cell energy minimization with neglect of thermal expansion might make it more difficult to ascertain and correct the flaws of the force field. It seems safer first to derive a set of parameters from ''constant cell'' calculations, and eventually to correct the force field for thermal effects.

The rms deviations between observed and minimum-energy heavy-atom coordinates of the five crystals, at various stages of the force-field fitting, are shown in Table IV. A first significant level of fitting is denoted as Fit-I in the first row of Table IV. It was reached by computing the electrostatic energy with the semiempirical (PM3) atomic charges, and adopting the effective value of 3 for the dielectric constant used for intramolecular interactions.

Then, substitution of the semiempirical charges on the sulfate group with the *ab initio* MP2 set yielded a decrease of the overall rmsd from 0.26 to 0.22 Å. This improvement is remarkable, considering that it was obtained without any adjustment of the force-field parameters and it concerns four of the five crystals.

Starting from Fit-I, with the MP2 charges, a new round of computations could only lower the overall rmsd to 0.21 Å. The modifications introduced in the force field at this level, besides some adjustments of bonded and nonbonded parameters, include the torsional term V_3 for the rotation of the SO_3^- group and the use of $\varepsilon = 1$, with scaled vicinal dipoles, for the computation of the intramolecular energy. The force field thus obtained was then tested against structure **6** of the Na^+ salt of 2-*N*-sulfated galactosamine (α-Gal-2NS Na). As shown in what follows, these computations indicated the need for a significantly more repulsive potential for the $S \cdots Na^+$ interaction. When the $S \cdots Na^+$ potential derived from **6** was applied to the three Na^+ salts of the *O*-sulfated sugars (**1**–**3**), the overall rmsd further decreased to 0.20 Å. This level of fitting, which here we regard as the final one, is denoted as Fit-II in the third row of Table IV.

In the fourth row of Table IV, we compare the results obtained with the same force field, but replacing the MP2 charges with those obtained at the SCF level, with the same basis set 6-31 + G**: the differences are rather small, considering that the SCF set presents a significantly larger polarization of the $S—O^-$ bonds (compare second and third columns of Table III). Only a further refinement of the force field could indicate a clear choice between the two sets of charges. As we shall discuss next, the major obstacle to improve the agreement with experiment is given by the present inability to understand the origin of the rearrangement of the hydrogen-bonding pattern occurring with the energy minimization of crystal **3**. Hence, it appears useless to pursue minor improvements.

In the second part of Table IV the geometrical features of the sulfate group computed for the structures of minimum energy are compared with the observed values. The rms deviation in the sulfate bond angles is 1°; the result is particularly satisfactory for bond angle C—O—S, whose observed values vary in a range of 6°, while the average discrepancy is only 0.7°. Also, the overall rmsd of the ring bond angles, shown in Table V, is rather small (1.6°). Considerably larger are the discrepancies concerning the dihedrals C—C—O—S and C—O—S—O^-; that is, although always within 10°, these deviations, originated by the difficulty to balance the strong interactions with mobile moieties (counterions and waters), correspond to significantly larger errors in the atomic positions of the sulfate group compared with the sugar ring. This is shown in Table VI; for example, in the case of crystal **1** the rmsd for the 12 C and O atoms of the ring is 0.07 Å, whereas for the remaining nonhydrogen atoms it is 0.30 Å.

HYDROGEN BONDING

Other structural details are presented in Tables VII and VIII, where we compare observed and computed H-bond lengths and short ion ⋯ ligand distances, respectively. Hydrogen bonding is well

TABLE V.

Comparison Between Experimental and Computed Ring Bond Angles in Crystals of *O*-Sulfated Monosaccharides (Fit-II, With Ab Initio MP2 Charges).

Crystal	α-MeGal-2S Na			α-MeGal-3S Na			α-MeGal-4S Na			α-MeGal-4S K			β-Glc-6S K			Overall	
bond angle	Obs.	Calcd.	Err.	Obs.	Calcd.	Err.	Obs.	Calcd.	Err.	Obs.	Calcd.	Err.	Obs.	Calcd.	Err.	rmsd	Av. dev.
C1—C2—C3	112.0	110.9	−1.1	109.5	111.7	2.2	112.4	11.06	−1.8	112.0	112.3	0.3	107.2	107.5	0.3	1.4	0.0
C2—C3—C4	109.8	109.9	0.1	112.0	108.1	−3.9	109.5	108.2	−1.3	109.6	110.0	0.4	112.3	110.4	−1.9	2.0	−1.3
C3—C4—C5	109.8	110.1	0.3	106.4	106.8	0.4	108.5	108.4	−0.1	109.6	108.3	1.3	111.5	112.9	1.4	0.9	0.7
C4—C5—O	109.7	110.3	0.6	107.1	110.1	3.0	107.8	109.9	2.1	108.1	109.5	1.4	110.5	111.2	0.7	1.8	1.6
C5—O—C1	113.8	114.3	0.5	113.5	113.0	−0.5	114.4	114.3	−0.1	114.6	114.0	−0.6	112.9	114.4	1.5	0.8	0.2
O—C1—C2	109.7	109.5	−0.2	111.9	111.6	−0.3	111.7	111.1	−0.6	110.7	111.3	0.6	108.1	109.4	1.3	0.7	0.2
O—C1—OG	112.0	113.5	1.5	111.6	112.8	1.2	111.6	112.4	0.8	111.2	111.7	0.5	107.0	105.3	−1.7	1.2	0.5
C2—C1—OG	108.5	110.4	1.9	107.5	108.6	1.1	108.2	109.6	1.4	109.3	111.7	2.4	111.3	112.9	1.6	1.7	1.7
C1—C2—O2	109.7	111.4	1.7	108.6	110.8	2.2	107.8	109.7	1.9	110.8	110.5	−0.3	111.0	111.8	0.8	1.5	1.3
C3—C2—O2	107.6	109.1	1.5	112.1	110.8	−1.3	111.2	110.6	−0.6	108.9	110.6	1.7	109.2	109.4	0.2	1.2	0.3
C2—C3—O3	109.1	110.1	1.0	104.1	108.8	4.7	110.3	112.9	2.6	110.8	112.3	1.5	109.5	110.6	1.1	2.6	2.2
C4—C3—O3	110.4	110.4	0.0	111.5	111.4	−0.1	112.1	112.0	−0.1	111.4	109.9	−1.5	109.7	109.2	−0.5	0.7	−0.4
C3—C4—O4	108.0	110.7	2.7	107.7	111.1	3.4	108.7	107.8	−0.9	108.0	109.9	1.9	109.9	109.2	−0.7	2.2	1.3
C5—C4—O4	111.1	110.7	−0.4	112.7	110.3	−2.4	106.8	108.1	1.3	107.4	107.2	−0.2	107.7	108.8	1.1	1.3	−0.1
C4—C5—C6	112.2	111.2	−1.0	117.3	114.5	−2.8	114.7	113.0	−1.7	115.8	111.8	−4.0	111.6	110.7	−0.9	2.4	−2.1
O—C5—C6	107.6	110.5	2.9	107.3	110.6	3.3	108.2	109.7	1.5	107.8	109.8	2.0	107.1	108.6	1.5	2.3	2.2
C5—C6—O6	113.2	113.0	−0.2	112.7	111.4	−1.3	108.4	110.1	1.7	109.9	110.1	0.2	108.5	109.6	1.1	1.1	0.3
C1—OG—C7	112.8	112.9	0.1	112.5	113.4	0.9	112.7	113.2	0.5	113.0	112.5	−0.5	—	—	—	0.6	0.2
rmsd			1.31			2.35			1.37			1.53			1.18	1.60	0.4

TABLE VI.

Discrepancies (Å) Between Observed and Calculated Atomic Positions in Six Crystals of Sulfated Monosaccharides (Fit-II, With *Ab Initio* MP2 Charges).

Crystal atom	α-MeGal-2S Na (1)	α-MeGal-3S Na (2)	α-MeGal-4S Na (3)	α-MeGal-4S K (4)	β-Glc-6S K (5)	α-Glc-2NS Na (6)
C1	0.031	0.134	0.170	0.068	0.133	0.257 (0.023)[a]
C2	0.083	0.082	0.284	0.095	0.127	0.274 (0.023)
C3	0.104	0.066	0.278	0.078	0.102	0.276 (0.023)
C4	0.082	0.174	0.169	0.072	0.090	0.223 (0.023)
C5	0.074	0.161	0.110	0.067	0.056	0.232 (0.023)
C6	0.047	0.125	0.118	0.132	0.084	0.168 (0.023)
C7	0.080	0.213	0.245	0.066	—	— (—)
O5	0.040	0.185	0.149	0.049	0.059	0.215 (0.109)
O2 / N2	0.185	0.099	0.496	0.149	0.229	0.310 (0.047)
O3	0.121	0.064	0.303	0.136	0.114	0.204 (0.141)
O4	0.063	0.174	0.083	0.069	0.129	0.238 (0.063)
O6	0.039	0.084	0.148	0.132	0.161	0.154 (0.074)
OG	0.060	0.136	0.147	0.070	0.166	0.354 (0.073)
Sx	0.292	0.092	0.165	0.132	0.051	0.306 (0.053)
Ox1	0.383	0.120	0.369	0.250	0.202	0.214 (0.092)
Ox2	0.296	0.163	0.068	0.373	0.235	0.283 (0.094)
Ox3	0.422	0.100	0.317	0.130	0.092	0.325 (0.010)
Ions	0.265	0.068	0.048/0.374	0.321	0.215	0.177 (0.039)
OW1	0.104	0.076	0.192	0.206	—	0.132 (0.174)
OW2	0.308	—	0.774	—	—	0.107 (0.118)
OW3	0.371	—	—	—	—	— (—)
rmsd	0.208	0.137	0.290	0.068	0.144	0.243 (0.082)

[a] Values in parentheses calculated after fitting by rototranslation.

reproduced in the two K^+ crystals. However, as we have already pointed out, in **4**, $O_3—H$ is bonded to O_5 rather than to O_{43}^-.

Also in crystal **1** bond $O_3—H$ rotates from the experimental orientation, but this time away from the H-bonded O_5; such a reorientation causes only a moderate lengthening of the $O_3 \cdots O_5$ distance and seems related to the optimization of the interaction with Na^+. This crystal, containing three water molecules per a.u., presents a complex hydrogen bond pattern, which is shown in Figure 2 and is mostly preserved during energy minimization, in spite of the mentioned displacements of the ionic groups. The major exception is the formation of a strong H-bond $O_{w2}—H_2 \cdots O_{22}^-$ at the expense of a weaker bond formed by a different crystallographic coupling of the same atoms.

The pattern of H bonds is also rather well reproduced in crystal **2**, where major deviations only concern the two distorted (bifurcated) bonds $O_2—H \cdots O_3$ and $O_6—H \cdots O_{33}^-$, without large changes in the distances between donors and acceptors.

Quite different is the case of α-MeGal-4S Na (**3**), where the original H-bond pattern is disrupted by energy minimization. As with crystal **1**, due to the presence of water molecules, the O atoms form a complex network of potential H bonds, shown in Figure 3. Although several patterns were investigated, only two of them correspond to energy minima: the one proposed from X-ray diffraction analyses[10,15] and the one obtained by (all-atom) energy minimization. Of the 7 H-bond-forming hydrogens contained in the a.u., $H—O_6$ and the two H_ws of water 2 maintain the initial orientation and H bonds ($O_6—H \cdots O_{w2}$, $O_{w2}—H_1 \cdots O_2$, $O_{w2}—H_1 \cdots O_G$, and $O_{w2}—H_2 \cdots O_6$). Also, $O_3—H$ remains bonded to O_5, although the weaker interaction $O_3—H \cdots O_4$ relaxes. On the contrary, water 1 is the moiety most affected by the rearrangement: bond $O_2—H$ rotates so as to break H bond $O_2—H \cdots O_{w1}$ and to form $O_2—H \cdots O_{41}^-$ (with a considerable shortening of the $O_2 \cdots O_{41}^-$ distance). This in turn breaks the $O_{w1}—H_1 \cdots O_{41}^-$ bond: water 1 rototranslates so

TABLE VII.
Comparison Between Observed and Calculated H-Bond Distances in Six Crystals of Sulfated Monosaccharides (Fit-II, with *Ab Initio* MP2 Charges).[a]

	α-MeGal-2s Na Observed		1 Calculated			α-MeGal-3s Na Observed		2 Calculated	
Donor—H ⋯ acceptor	$R_{D\cdots A}$	($R_{H\cdots A}$)	$R_{D\cdots A}$	($R_{H\cdots A}$)	Donor—H ⋯ acceptor	$R_{D\cdots A}$	($R_{H\cdots A}$)	$R_{D\cdots A}$	($R_{H\cdots A}$)
	αMeGal-2S NA		1			α-MeGal-3S Na		2	
O_3—H ⋯ O_5	2.753	(2.098)	2.817	(lost)	O_2—H ⋯ O_3	2.977	(2.432)	2.883	(lost)
O_4—H ⋯ O_{21}	2.702	(1.904)	2.730	(1.937)	O_2—H ⋯ O_4	2.898	(1.999)	2.906	(2.043)
O_6—H ⋯ W_3	2.721	(1.840)	2.783	(1.841)	O_4—H ⋯ O_{32}	2.838	(1.937)	2.826	(1.911)
W_1—H_1 ⋯ O_{21}	2.914	(1.980)	2.942	(2.011)	O_6—H ⋯ O_{31}	2.925	(2.040)	3.038	(2.185)
W_1—H_2 ⋯ O_{22}	3.222	(2.604)	3.400	(lost)	O_6—H ⋯ O_{33}	3.244	(new)	3.225	(2.546)
W_1—H_2 ⋯ O_{22}	2.887	(new)	2.730	(1.908)	W_1—H_1 ⋯ O_6	3.229	(2.359)	3.215	(2.328)
W_2—H_2 ⋯ $O_{22'}$	2.785	(1.884)	2.756	(1.925)	W_1—H_2 ⋯ O_{22}	3.326	(2.404)	3.323	(2.423)
W_2—H_1 ⋯ O_4	2.877	(2.304)	2.827	(2.074)	W_1—H_2 ⋯ O_{22}	3.326	(2.404)	3.323	(2.423)
W_3—H_1 ⋯ O_2	3.177	(2.242)	3.261	(2.401)					
W_3—H_1 ⋯ O_{22}	3.053	(2.453)	2.947	(2.270)					
W_3—H_2 ⋯ O_6	2.827	(1.908)	2.910	(1.991)					
	αMeGal-4S Na		3			α-MeGal-4S K		4	
O_2—H ⋯ W_1	2.783	(1.920)	2.728	(lost)	O_2—H ⋯ O_{41}	2.772	(1.893)	2.791	(1.965)
O_2—H ⋯ O_{41}	3.462	(new)	2.729	(1.872)	O_3—H ⋯ O_5	2.825	(1.939)	2.813	(1.976)
O_3—H ⋯ O_5	2.775	(1.871)	2.803	(2.034)	O_6—H ⋯ W_1	2.662	(1.707)	2.708	(1.764)
O_6—H ⋯ W_2	2.672	(1.721)	2.746	(1.810)	W_1—H_1 ⋯ O_6	2.787	(1.835)	2.823	(1.874)
W_1—H_1 ⋯ O_{41}	2.751	(1.912)	3.055	(lost)	W_1—H_2 ⋯ O_G	2.970	(2.158)	2.828	(1.949)
W_1—H_1 ⋯ O_{43}	2.828	(new)	2.788	(2.043)	W_1—H_2 ⋯ O_2	3.107	(2.420)	3.029	(2.426)
W_1—H_2 ⋯ O_{43}	2.828	(2.027)	2.788	(lost)					
W_1—H_2 ⋯ O_2	2.783	(new)	2.728	(1.878)					
W_1—H_2 ⋯ O_3	3.303	(new)	2.866	(2.259)					
W_3—H_1 ⋯ O_G	2.832	(2.304)	2.898	(2.091)		α Gic-2NS Na		6	
W_3—H_1 ⋯ O_2	3.135	(2.320)	2.902	(2.135)					
W_3—H_2 ⋯ O_6	2.896	(1.950)	2.881	(1.938)	N_2—H ⋯ O_{22}	3.015	(2.007)	3.033	(2.011)
					O_3—H ⋯ O_{23}	2.786	(1.902)	2.760	(1.875)
	β Gic-6S K		5		O_4—H ⋯ O_{22}	2.758	(1.884)	2.866	(1.958)
					O_6—H ⋯ N_2	3.109	(2.526)	3.068	(2.510)
O_2—H ⋯ O_{61}	2.847	(1.942)	2.880	(2.065)	O_G—H ⋯ W_2	2.984	(2.299)	2.889	(2.166)
O_3—H ⋯ O_2	2.872	(1.998)	2.833	(2.002)	W_1—H_1 ⋯ O_{23}	2.672	(1.846)	2.624	(1.893)
O_4—H ⋯ O_3	2.687	(1.855)	2.695	(1.960)	W_1—H_2 ⋯ O_45	2.753	(1.845)	2.802	(1.873)
O_4—H ⋯ O_2	3.240	(2.485)	3.030	(2.227)	W_2—H_1 ⋯ O_{32}	2.862	(1.916)	2.836	(1.892)
O_G—H ⋯ O_{61}	2.832	(1.876)	2.847	(1.997)	W_2—H_2 ⋯ O_{21}	3.019	(2.106)	2.988	(2.056)

[a] Observed distances $R_{H\cdots A}$ calculated after optimization of H coordinates.

that O_{w1}—H_1 replaces O_{w1}—H_2 in the bond O_{w1}—H ⋯ O_{43}^- and O_{w1}—H_2 forms a bifurcated H bond with O_3 and O_2.

Such a complex rearrangement is related to a subtle interplay of strong electrostatic interactions, which the present model appears unable to balance perfectly. In the intermediate stage of the force-field fitting (prior to Fit-I), the structure **3** of minimum energy showed excellent agreement with experiment (rmsd of 0.15 Å). Then new modifications of the force field led to the different H-bond pattern (and rmsd of ~ 0.30 Å). Subsequent computations showed that, also with the previous force field parameters, the alternative pattern is always the more stable one, the energy difference being ~ 5 kcal mol^{-1} per a.u. With the final parameters (Fit-II), the same minimized structure is found independently from the starting point. Quite similar results were also obtained when the unit cell parameters and the fractional coordinates deter-

TABLE VIII.

Comparison Between Observed and Calculated Ion ⋯ Ligand Distances (Å) in Six Crystals of Sulfated Monosaccharides (Fit-II, with Ab Initio MP2 Charges).

Crystal	α-MeGal-2S Na	1		α-MeGla-3S Na	2
Ion ⋯ ligand	$R_{obs.}$	$R_{calcd.}$	Ion ⋯ ligand	$R_{obs.}$	$R_{calcd.}$
$Na^+ \cdots W_1$	2.313	2.261	$Na^+ \cdots W_1$	2.342	2.301
$Na^+ \cdots W_2$	2.440	2.363	$Na^+ \cdots O_2$	2.330	2.360
$Na^+ \cdots W'_2$	2.369	2.359	$Na^+ \cdots O_6$	2.445	2.449
$Na^+ \cdots O_3$	2.320	2.293	$Na^+ \cdots O_5$	2.458	2.560
$Na^+ \cdots O_4$	2.499	2.373	$Na^+ \cdots O_{33}$	2.618	2.639
$Na^+ \cdots O_{23}$	2.662	2.769	$Na^+ \cdots O_G$	2.692	2.624
			$Na^+ \cdots O_4$	2.814	2.547
Crystal[a]	α-MeGal-4S Na	3		α-MeGal-4S K	4
Ion ⋯ ligand	$R_{obs.}$	$R_{calcd.}$	Ion ⋯ ligand	$R_{obs.}$	$R_{calcd.}$
$Na_1^+ \cdots O_{42}$	2.359	2.365	$K^+ \cdots O_{42}$	2.680	2.731
$Na_1^+ \cdots O_3$	2.404	2.498	$K^+ \cdots O'_{42}$	2.703	2.675
$Na_1^+ \cdots O_{43}$	2.416	2.588	$K^+ \cdots O_{43}$	2.681	2.768
$Na_2^+ \cdots W_1$	2.338	2.237	$K^+ \cdots O_3$	2.882	2.943
$Na_2^+ \cdots O_{42}$	2.391	2.522	$K^+ \cdots O_{43}$	2.970	2.722
$Na_2^+ \cdots O_{41}$	2.650	2.922	$K^+ \cdots O_{41}$	3.012	2.992
			$K^+ \cdots O_2$	3.059	3.002
Crystal	β-Glc-6S K	5		α-Glc-2NS Na	6
Ion ⋯ ligand	$R_{obs.}$	$R_{calcd.}$	Ion ⋯ ligand	$R_{obs.}$	$R_{calcd.}$
$K^+ \cdots O_{62}$	2.735	2.781	$Na^+ \cdots W_1$	2.362	2.288
$K^+ \cdots O'_{62}$	2.792	2.747	$Na^+ \cdots O_6$	2.337	2.307
$K^+ \cdots O_{63}$	2.715	2.727	$Na^+ \cdots O_{21}$	2.382	2.388
$K^+ \cdots O'_{63}$	2.760	2.787	$Na^+ \cdots O_3$	2.386	2.352
$K^+ \cdots O_5$	2.760	2.857	$Na^+ \cdots W_2$	2.407	2.296
$K^+ \cdots O_G$	2.796	2.765	$Na^+ \cdots O_5$	2.427	2.549

[a] In crystal **3**, twofold axes generate equal $Na^+ \cdots$ ligand distances.

mined at low temperature by Kanters et al.[15] were utilized instead of the data by Lamba et al.[10] So far, it has not been possible to identify a single flaw in the force field responsible for the overstabilization of a hypothetical structure which definitely does not correspond to the observed one.

Although water 1 shows the largest displacement (nearly 0.8 Å), its interactions may be not the driving force of the rearrangement. Analysis of the energy contributions indicated that (after all-atom optimization) the net balance favors the groups $-C_2-O-H$ and $C_4-O-SO_3^- \cdot Na^+$; on the other hand, if the final H-bond pattern is provided in the starting point of a run where the heavy atoms are fixed at their observed positions, the computed structure is more stable than the original one by 1.2 kcal mol^{-1}, the alternative pattern mainly favoring the water 1 molecule (bending and van der Waals terms). At present, we suspect that the current lone-pair model, as defined in our force field for water as well as for alcoholic —OH, may overestimate the role of the lone pairs in the calculation of the electrostatic energy. We leave the goal of defining whether the residual discrepancies pointed out in this section arise from the said factor to future refinements based on a larger number of structures.

α-GAL-2NS NA

The purpose of the calculations on the crystal of α-Gal-2NS Na (**6**) was to test the force field derived from the *O*-sulfated monosaccharides, rather than to extend the force-field fitting to *N*-sulfated compounds, considering that, for the latter purpose, one crystal structure is not sufficient. Thus,

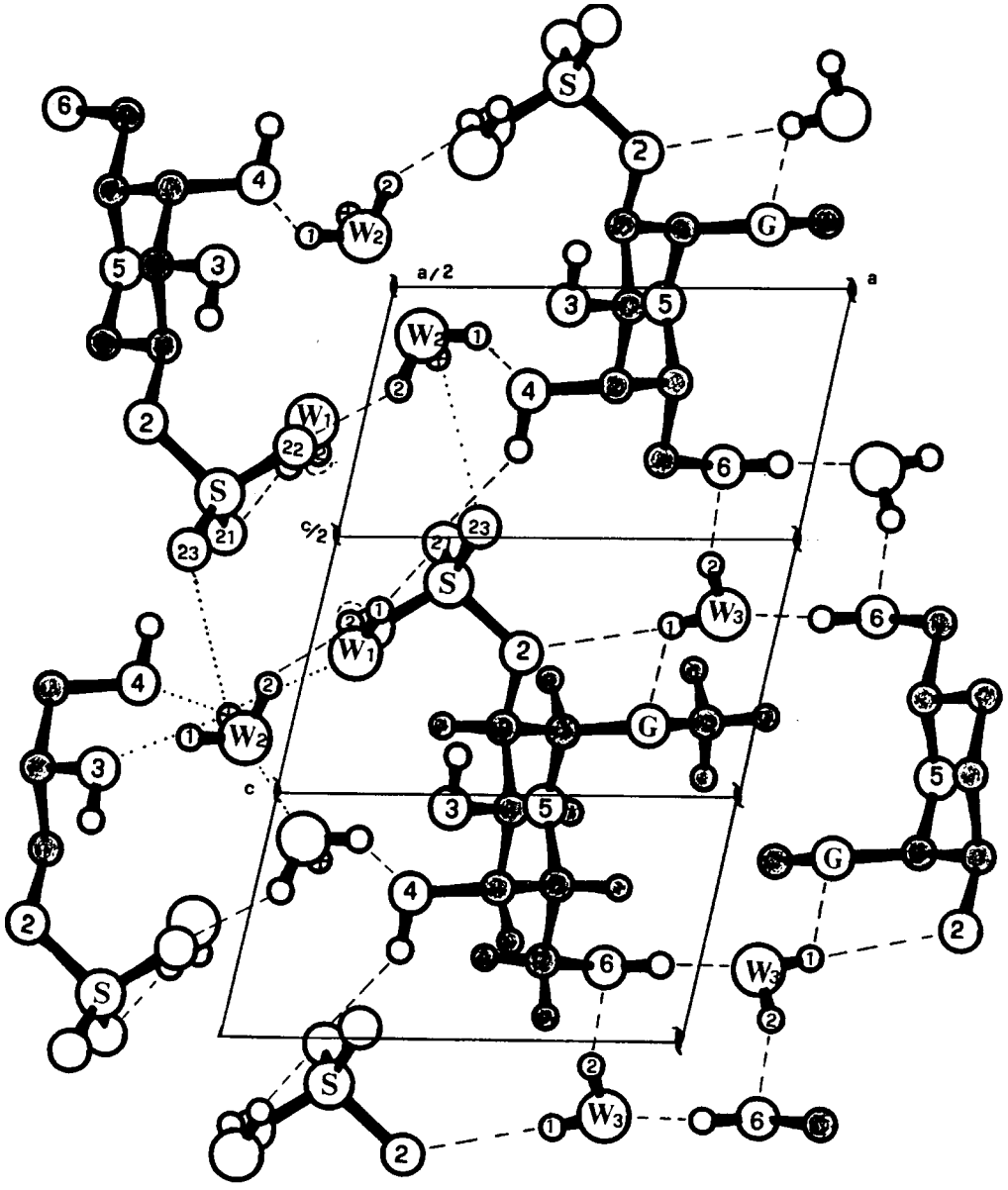

FIGURE 2. View normal to the *b*-axis of the minimum-energy structure of crystal **1** (α-MeGal-2S Na). Dashed lines show hydrogen bonds, dotted lines indicate ion coordination.

while it was necessary to try tentative parameters for the $N \cdots Na^+$ interaction, and, by analogy with the treatment of the *O*-sulfated compounds, the atomic charges obtained from *ab initio* computations on *N*-methylsulfate[22,26] (see Table III, fourth column) were adopted for the $-C-NH-SO_3^-$ group, for the rest we first utilized the previous (bonded and nonbonded) parameters[3] concerning the N atom. This includes the assumption of a lone-pair pseudo-atom on N.

Owing to the sp^3 nature of the sulfamido nitrogen, as already observed in the model compounds[29] utilized for the old parameterization[3] and confirmed by the *ab initio* studies,[22,26] the H_N atom may occupy two different positions, corresponding to alternative ''configurations'' of the N atom. Hence, two starting points, after hydrogen and lone-pair optimization, were considered for the ''all-atom'' energy minimization of **6**. It was found that, while the minimum-energy structure

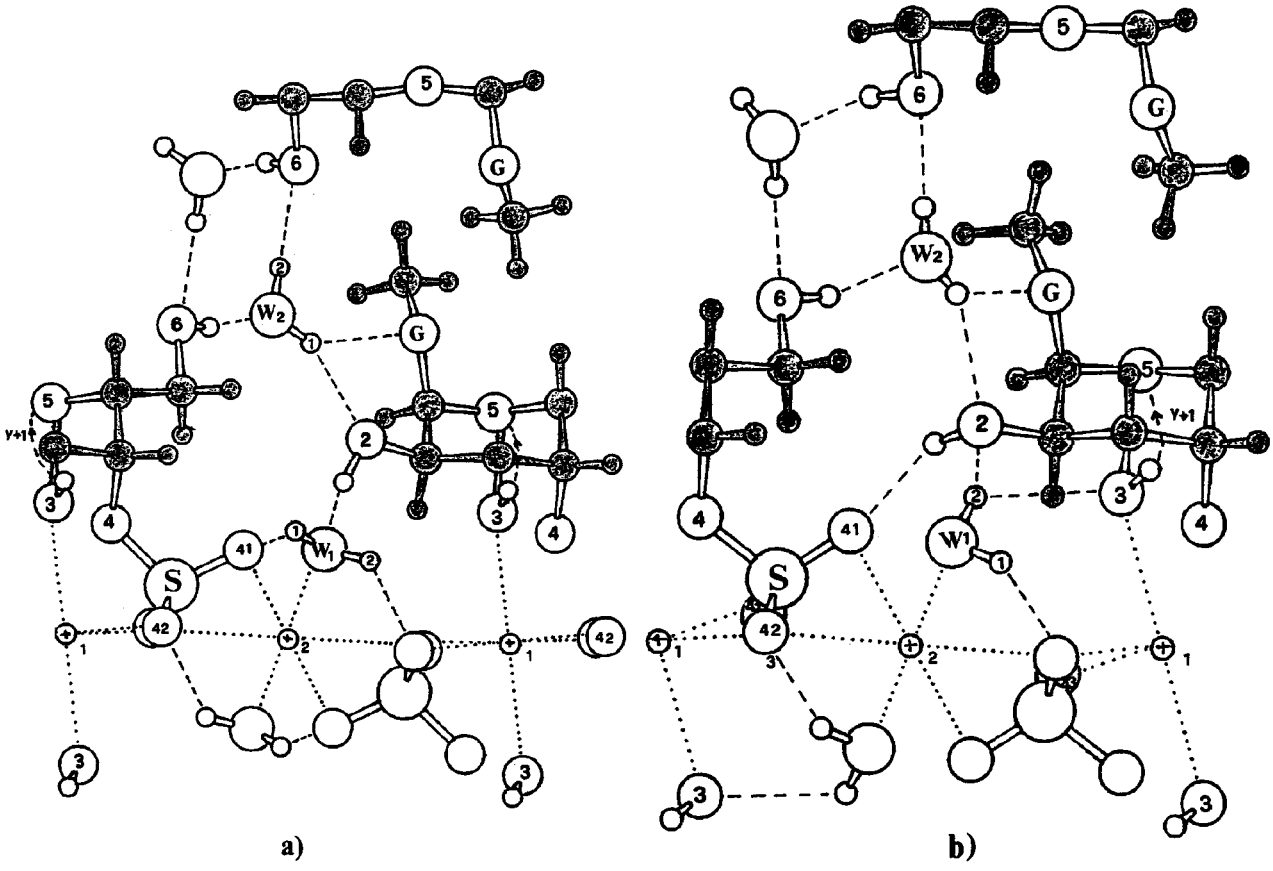

FIGURE 3. View normal to the *b*-axis of crystal **3** (α-MeGal-4S Na). (a) Experimental structure; (b) energy minimum.

reached from one N configuration is in reasonable agreement with experiment (rms displacement of heavy atoms of 0.26 Å), in the other case energy minimization leads to a structure, more stable than the former, quite different from the experimental one (rmsd of 0.52 Å), where atom N_2 replaces O_5 as one of the six Na^+ ligands and the conformation of the sulfamino group deviates by as much as 60°. Analysis of distances and energy contributions indicated that two factors were required to avoid such an artifact: (a) to remove the lone pair from N (actually its bonded interactions were retained to keep the previous N parameters); and (b) to use a more repulsive potential for the $S \cdots Na^+$ interaction. With such modifications (the $S \cdots Na^+$ potential, afterwards used for the final calculations on crystals **1**–**5**, is given in Table II) the alternative structure is less stable by 3.5 kcal · mol^{-1}, while the rms deviation for the former structure decreases to 0.243 Å. Very similar results were obtained by replacing the *ab initio* MP2 charges on $C—NH—SO_3^-$ by the SCF set (rmsd of 0.242 Å). The agreement with experiment is somewhat worse than the average value for the five *O*-sulfated crystals. However, the error is uniformly distributed over all atoms, as it is largely due to a rigid rotation of the asymmetric unit; in fact, if the computed structure of the asymmetric unit is rototranslated to best fit the observed atomic positions, the rms deviation (representing conformational differences) is reduced to 0.08 Å (only 0.06 Å if water molecules are not included). This is confirmed by the good agreement between computed and observed features, shown in Tables IX and X.

Conclusions

This work provides an improved force field for sulfated carbohydrates, which yields an overall discrepancy of 0.21 Å between the atomic positions computed for the structures of minimum energy (at constant cell) and those observed by X-ray diffraction in six crystals of Na^+ and K^+ salts. When calculated and experimental structures are

TABLE IX.
Results of Energy Minimizations on α-Glc-2NS Na Crystal: rms Deviation Between Computed and Experimental Atomic Coordinates (Nonhydrogen Atoms), and Comparison Between Computed nd Experimental Geometrical Features of Sulfate Group (Fit-II, With Ab Initio MP2 Charges; see text).

CCrystal × *n* H_2O	α-Glc-2NS Na		
rms dev. (Å)	2		
	0.243		
	Obs.	Calcd.	Err.
Bond length			
N—S	1.639	1.639	.000
S—$O_{g/g'}$	1.452	1.443	−.009
S—O_a	1.466	1.439	−.027
Bond angle			
C—N—S	118.0	118.7	0.7
N—S—O_g	111.3	107.3	−4.0
N—S—$O_{g'}$	105.9	105.1	−0.8
N—S—O_a	102.1	102.8	0.7
O_g—S—$O_{g'}$	112.0	113.0	1.0
O_a—S—O_g	111.6	113.8	2.2
O_a—S—$O_{g'}$	113.4	113.7	0.3
Dihedral			
C—C—N—S	−124.9	−124.8	0.1
C—N—S—O_a	−170.6	−174.0	−3.4

TABLE X.
Comparison Between Experimental and Computed Ring Bond Angles in α-Glc-2NS Na Crystal (6). (Fit-II, With Ab Initio MP2 Charges).

Crystal	α-Glc-2NS Na		
bond angle	Obs.	Calcd.	Err.
C1—C2—C3	111.7	111.1	−0.6
C2—C3—C4	110.0	109.8	−0.2
C3—C4—C5	110.6	111.6	1.0
C4—C5—O	110.6	112.2	1.6
C5—O—C1	113.1	114.9	1.8
O—C1—C2	110.9	113.7	2.8
O—C1—OG	112.2	109.7	−2.5
C2—C1—OG	109.9	111.8	1.9
C1—C2—N2	109.9	107.4	−2.5
C3—C2—N2	111.4	112.2	0.8
C2—C3—O3	108.3	110.9	2.6
C4—C3—O3	111.4	108.8	−2.6
C3—C4—O4	108.6	109.0	0.4
C5—C4—O4	108.7	109.1	0.4
C4—C5—C6	113.6	111.3	−2.5
O—C5—C6	106.3	109.1	2.8
C5—C6—O6	111.0	109.9	−1.1
rmsd			1.89

compared after best-fitting the models by roto-translation, the overall discrepancy is 0.11 Å, showing that the internal geometry of the sulfated monosaccharides is rather well reproduced by the calculations.

The novel parameters of the force field, which is otherwise based on MM2(87), have been derived *empirically* from the crystal structure analysis, but the use of the *ab initio* charge distribution for the sulfate groups has been decisive for improving the quality of the force-field fitting. The agreement with experiment is far from being homogeneous both in the different crystals examined and in the various molecular domains. The treatment of strong electrostatic interactions should be improved; in particular, the current use of point charges for representing lone pairs and the neglect of polarization effects could be questioned.

Accurate *ab initio* quantum-mechanical computations on complex fragments of the microcrystals could help in the refinement of the force field, especially with regard to evaluating the relative stability of nonobserved structures. In other words, such computations may play the role of extra ''experiments,'' beyond the single observation of the stable structure. Moreover, Withfield and Tang[19] and Huige and Altona[20] have recently shown, at least partially, that the application of *ab initio* methods to model compounds may directly provide parameters to be incorporated in the force field. Unfortunately, for various practical reasons, we cannot readily test the general validity of their proposals against the crystal structures investigated here. For one of these crystals, however, a comparison is possible, because Huige and Altona[20] applied their extensions of the CHARMm force fields, derived from *ab initio* calculations on sulfate and sulfamate model compounds, to compute the minimum energy structure of compound **5** (the K^+ salt of β-D-6-sulfoglucopyranose), starting from the same experimental structure[9] utilized in the present work. While most geometrical features are reproduced with similar accuracy in their and in our calculations, it is not surprising that the rms deviation of the heavy atoms, computed at constant lattice dimensions, is significantly larger in their case (cf. the values of 0.305 Å with CHARMm32 and 0.350 Å with CHARMm33 in Table XVI of ref. 20, with the values 0.144–0.201 Å in Table IV of this work), as no parameterization of the intermolecular interactions was adjusted to compound **5**. On the other hand, they find a very good agreement with experiment (particularly with

CHARMm33) if the minimum-energy structure is computed *in vacuo*, while a significant improvement is observed when the lattice constants are relaxed (rms deviation of 0.189 Å), indicating the need for some corrections in the treatment of nonbonded interactions.[20] Thus, we may conclude that the best way for defining a highly reliable force field may consist in coupling accurate quantum-mechanical computations with a crystal structure analysis of the type attempted in the present work.

Acknowledgments

We are grateful to Drs. Doriano Lamba and William Mackie for providing us with the crystal coordinates prior to publication. This work was partially supported by Progetto Finalizzato Chimica Fine of the Italian CNR through a 2-year fellowship to P.P.

Appendix

α-MeGal-2S Na = methyl-2-*O*-sulfo-α-D-galactopyranoside—$Na^+ \cdot H_2O$

α-MeGal-3S Na = methyl-3-*O*-sulfo-α-D-galactopyranoside—$Na^+ \cdot H_2O$

α-MeGal-4S Na = methyl-4-*O*-sulfo-α-D-galactopyranoside—$Na^+ \cdot 2\ H_2O$

α-MeGal-4S K = methyl-4-*O*-sulfo-α-D-galactopyranoside—$K^+ \cdot H_2O$

β-Glc-6*S* K = 6-*O*-sulfo-*β*-D-glucopyranose—K^+;

α-Gal-2NS Na = 2-deoxy, 2-sulfamino-α-D-glucopyranose—$Na^+ \cdot 2\ H_2O$

References

1. R. L. Jackson, S. J. Bush, and A. D. Cardin, *Physiol. Rev.* **71**, 481–525 (1991).
2. M. Ragazzi, D. R. Ferro, and A. Provasoli, *J. Comput. Chem.*, **7**, 105–112 (1986).
3. D. R. Ferro, A. Provasoli, M. Ragazzi, G. Torri, B. Casu, G. Gatti, J.-C. Jaquinet, P. Sinaÿ, M. Petitou, and J. Choay, *J. Am. Chem. Soc.*, **108**, 6773–6778 (1986).
4. M. Ragazzi, D. R. Ferro, B. Perly, P. Sinaÿ, M. Petitou, and J. Choay, *Carbohydr. Res.* **195**, 169–185 (1990).
5. M. Ragazzi, A. Provasoli, and D. R. Ferro, *ACS Symp. Ser.*, **430**, 332–344 (1990).
6. D. Lamba, A. L. Segre, M. Ragazzi, D. R. Ferro, and R. Toffanin, *Carbohydr. Res.*, **209**, c13–c15 (1991).
7. D. R. Ferro, J. Gajdoš, M. Ragazzi, F. Ungarelli, and S. Piani, *Carbohydr. Res.*, **277**, 25–38 (1995).
8. Y. Nawata, K. Ochi, M. Shiba, K. Morita, and Y. Iitaka, *Acta Cryst.*, **B37**, 246–249 (1981).
9. D. Lamba, W. Mackie, B. Sheldrick, P. Belton, and S. Tanner, *Carbohydr. Res.*, **180**, 183–193 (1988).
10. D. Lamba, S. Glover, W. Mackie, A. Rashid, B. Sheldrick, and S. Pérez, *Glycobiology*, **4**, 151–163 (1994).
11. E. A. Yates, W. Mackie, and D. Lamba, *Int. J. Biol. Macromol.*, **17**, 219–226 (1995).
12. D. Lamba, W. Mackie, A. Rashid, B. Sheldrick, and E. A. Yates, *Carbohydr. Res.*, **241**, 89–98 (1993).
13. W. Mackie, E. A. Yates, and D. Lamba, *Carbohydr. Res.*, **266**, 65–74 (1995).
14. R. Vega, A. López-Castro, and R. Márquez, *Acta Cryst.* **C42**, 1066–1068 (1986).
15. J. A. Kanters, B. van Dijk, and J. Kroon, *Carbohydr. Res.*, **212**, 1–11 (1991).
16. J. G. Fernández-Bolaños, J. Morales, S. García, M. J. Diánez, M. D. Estrada, A. López-Castro, and S. Pérez, *Carbohydr. Res.*, **248**, 1–14 (1993).
17. A. J. Polvorinos, R. R. Contreras, D. Martin-Ramos, J. Romero, and M. A. Hidalgo, *Carbohydr. Res.*, **257**, 1–10 (1994).
18. W. H. Ojala, K. E. Albers, and W. B. Gleason, *Poster B17, ACA meeting*, Atlanta, GA (1994).
19. D. M. Whitfield and T.-H. Tang, *J. Am. Chem. Soc.*, **115**, 9648–9654 (1993).
20. C. J. M. Huige and C. Altona, *J. Comput. Chem.*, **16**, 56–79 (1995).
21. (*a*) N. L. Allinger and Y. H. Yu, *QCPE*, **13**, 395 (1980); (*b*) N. L. Allinger, R. A. Kok, and M. R. Iman, *J. Comput. Chem.*, **9**, 591–595 (1988).
22. D. R. Ferro, P. Pumilia, A. Cassinari, and M. Ragazzi, *Int. J. Biol. Macromol.*, **17**, 131–136 (1995).
23. B. H. Besler, K. M. Merz Jr., and P. A. Kollman, *J. Comp. Chem.* **11**, 431 (1990).
24. J. J. P. Stewart, *J. Comput.-Aid. Mol. Des.*, **4**, 1 (1990).
25. M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Robb, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, *Gaussian 92 (Rev. E3)*, 1992, Gaussian Inc., Pittsburgh, PA.
26. M. Ragazzi and D. R. Ferro, *Theochem. special issue* (in press).
27. D. R. Ferro and M. Ragazzi, *Molecular mechanics program CHAMP* (**C**onformational [**H**yper] **A**nalysis **M**ilan **P**ackage), ICM-CNR, Milano (1993).
28. D. R. Ferro, S. Brückner, S. V. Meille, and M. Ragazzi, *Macromolecules*, **23**, 1676–1680 (1990); *Macromolecules*, **24**, 1156–1160 (1991); *Macromolecules*, **25**, 5231–5235 (1992).
29. A. J. Morris, C. H. L. Kennard, J. R. Hall, and G. Smith, *Inorg. Chim. Acta*, **62**, 247–252 (1982).